

Figure 1 - Error Recovery Architecture

20949P-000800US George B. F. Yee, Reg. No. 37,478 Telephone: 650-326-2400 Inventors: Richard Meyer et al. Title: Failover Processing in a Storage System Sheets of drawings 2 or 36

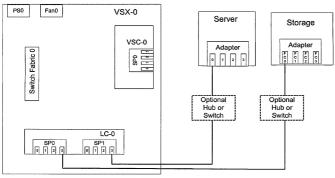


Figure 2 - Non-Fault Tolerant Configuration

20949P-000800US George B. F. Yee, Reg. No. 37,478 Telephone: 650-326-2400 Inventors: Richard Meyer et al. Title: Failover Processing in a Storage System Sheets of drawings 3 of 36

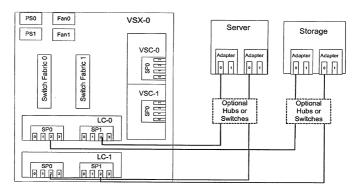


Figure 3 - Fault Tolerant Configuration

20949P-000800US George B. F. Yee, Reg. No. 37,478 Telephone: 650-326-2400 Inventors: Richard Meyer et al. Trite: Failover Processing in a Storage System Sheets of drawings 4 of 36

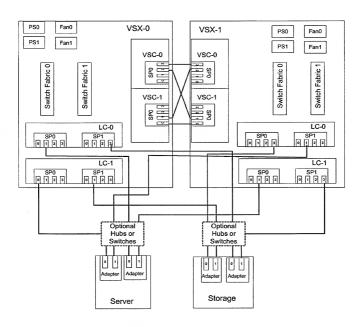


Figure 4 - High Availability Configuration

20949P-000800US George B. F. Yee, Reg. No. 37,478 Telephone: 509-326-2400 Inventors: Richard Meyer et al. Title: Pailover Processing in a Storage System Shoets of drawings 5 of 36

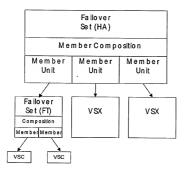


Figure 5 Components of a Failover Set

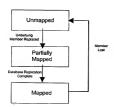


Figure 6: Member Unit State Diagram

20949P-000800US George B. F. Yee, Reg. No. 37,478 Telephone: 650-326-2400 Inventors: Riehard Meyer et al. Title: Failover Processing in a Storage System Sheets of drawings 6 of 36

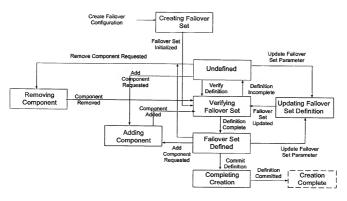


Figure 7 - Creating a Failover Set

20949P-000800US George B F. Yee, Reg. No. 37,478 Telephone: 650-326-2400 Inventors: Richard Meyer et al. Trite: Pailover Processing in a Storage System Sheets of drawings 7 of 36

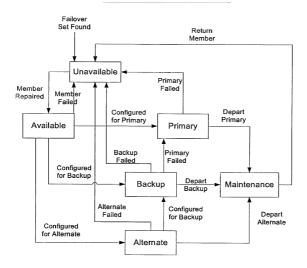


Figure 8 - Member State Diagram

20949P-000800US George B F. Yee, Reg. No. 37,478 Telephone: 650-326-2400 Inventors: Richard Meyer et al. Title: Failover Processing in a Storage System Sheets of drawings 8 of 36

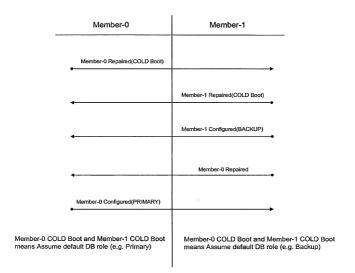


Figure 9 - Member Arbitration for COLD Boot

20949P-000800US George B. F. Yee, Reg. No. 37,478 Telephone: 650-326-2400 Inventors: Richard Meyer et al. Title: Fatlover Processing in a Storage System Shects of drawings 9 0 136

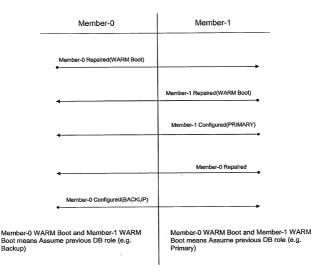


Figure 10 - Member Arbitration for WARM Boot

20949P-000800US George B. F. Yee, Reg. No. 37,478 Telephone: 650-326-2400 Inventors: Richard Meyer et al. Title: Failover Processing in a Storage System Sheets of drawings 10 of 36

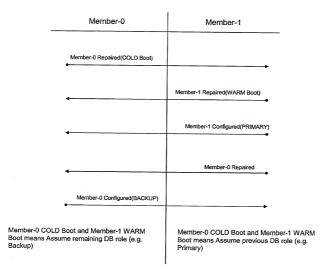


Figure 11 - Member Arbitration for Mixed Boot

	Event					
Old State	Mi Repaired	Mj Repaired	Mi Configured	Mi Configured Mj Configured Mi Failed	Mi Failed	Mj Failed
 {Mi,Mj} Unavail, {} Avail, {} Primary, {} Backup* 	New St: 3	New St: 2			New St: 1	New St; 1
	Action: A	Action: B			Action: S	
2. (Mi) Unavail, (Mj) Avail, () Primary, () Backup	New St: 4			New St: 8	New St: 2	
	Action: C			Action: D	Action: S	
3. (Mj) Unavail, (Mi) Avail, () Primary, () Backup		New St: 4	New St: 9			New St: 3
		Action: E	Action: F			Action: T
4. {} Unavail, {Mi, Mj} Avail, {} Primary, {} Backup			New State: 7	New St: 6		
			Action: G	Action: H		
5a. {} Unavail, {} Avail, {Mi} Pri, {Mj} Backup					New St: 8	New St: 9
					Action: I	Action: J
5b. {} Unavail, {} Avail, {Mj} Pri, {Mi} Backup					New St: 8	New St: 9
					Action: I	Action: J
6. {} Unavail, {Mi} Avail, {Mj} Pri, {} Backup	New St: 6		New St: 5a, 5b			New St: 3
	Action: K		Action: L			Action: M
7. {} Unavail, {Mj} Avail, {Mi} Pri, {} Backup		New St: 7		New St: 5a, 5b	New St: 2	
		Action: N		Action: O	Action: P	
8. (Mi) Unavail, () Avail, (Mj) Pri, {} Backup	New St: 6					New St: 1
	Action: C					Action: Q
9. {Mj} Unavail, {} Avail, {Mi} Pri, {} Backup		New St: 7			New St: 1	
		Action: E			Action: R	
* Initial State						

20949P-000800US George B. F. Yee, Reg. No. 37,478 Telephone: 650-326-2400 Inventors: Richard Meyer et al. Title, Failover Processing in a Storage System Sheets of drawings 11 of 36

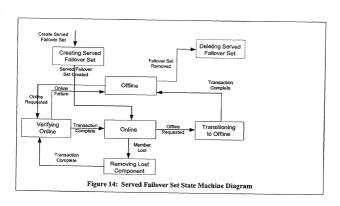
Figure 12 - 2 Member State Table

20949P-000800US George B. F. Yee, Reg. No. 37,478 Telephone: 650-326-2400 Inventors: Richard Meyer et al. Title: Failover Processing in a Storage System Sheets of drawings 12 of 36

Action Routines	Description					
1	1. Send "Mi repaired" to Mj, if Mj is not failed. 2. Set timer to send "Mi repaired" to Mi					
2	1. Send "Mj repaired" to Mi, if Mi is not failed. 2. Set timer to send "Mj repaired" to Mj					
A	1. If Mi and configured send "Mi configured" to Mj. 2. Set timer to send "Mi configured" to Mi. 3.					
В	1. If Mj and configured send "Mj configured" to Mi. 2. Set timer to send "Mj configured" to Mj. 3.					
С	1. If Mj, echo event back to Mi. 2. If Mi and configured send "Mi configured" to Mj. 3. Set timer to					
D	1. If Mj become Primary. 2. Otherwise, nop.					
E	1. If Mi, echo event back to Mj. 2. If Mj and configured send "Mj configured" to Mi. 3. Set timer to					
F	1. If Mi become Primary. 2. Otherwise, nop.					
G	1. If Mi become Primary. 2. Otherwise, echo event back to Mi.					
H	1. If Mj become Primary. 2. Otherwise, echo event back to Mj.					
I	If Mj become Primary. 2. If Mi become Backup.					
J	If Mi become Primary. 2. If Mj become Backup.					
K	1. If Mj echo event back to Mi. 2. Otherwise, nop					
L	1. If Mj determine Member Role. 2. Send "Mi configured" to Mi when done. 3. If Mi determine					
M	1. If Mj perform Fail-Stop processing. 2. Send "Mj Failed" to Mi. 3. Otherwise become Primary after					
N	1. If Mi echo event back to Mj. 2. Otherwise, nop					
0	1. If Mi determine Member role. 2. Send "Mj configured" to Mj when done. 3. If Mj determine					
P	1. If Mi perform Fail-Stop processing. 2. Send "Mi Failed" to Mj. 3. Otherwise become Primary after					
Q	If Mj perform Fail-Stop processing for Mj 2. Otherwise nop.					
R	1. If Mi perform Fail-Stop processing for Mi. 2. Otherwise nop.					
S	1. Perform Fail-Stop processing for Mi					
T	1. Perform Fail-Stop processing for Mj					

Figure 13 -Action Routines for a 2 Node Configuration

20949P-000800US George B. F. Yee, Reg. No. 37,478 Telephone: 650-326-2400 Inventors: Richard Meyer et al. Title: Failover Processing in a Storage System Sheets of drawings 13 of 36



20949P-000800US George B. F. Yee, Reg. No. 37,478 Telephone: 650-326-2400 Inventors: Richard Meyer et al. Title: Failover Processing in a Storage System Sheets of drawings 14 of 36

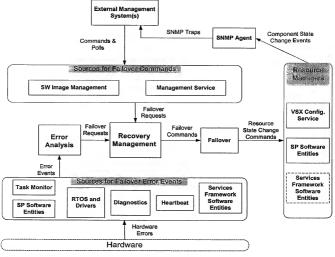
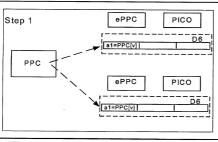
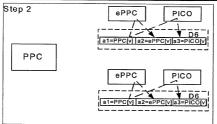


Figure 15 - Fault Detection and Analysis Architecture

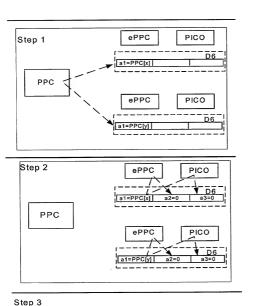




Step 3

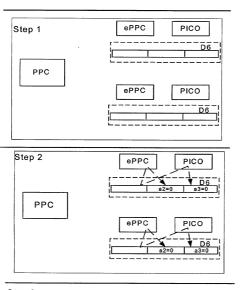
majority(a1,a2,a3) = majority(v,v,v) = v, No faults

Figure 16 - No faults



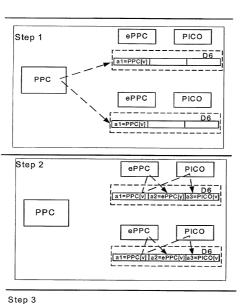
majority(a1,a2,a3) = majority(x,0,0) = 0, transmitter fault

Figure 17 - Transmitter fault (sends a bad value)



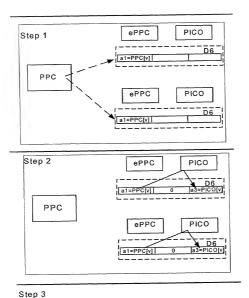
Step 3
majority(a1,a2,a3) = majority(0,0,0) = 0, transmitter fault

Figure 18 - Transmitter fault (doesn't send a value)



majority(a1,a2,a3) = majority(v,y,v) = v, Receiver fault

Figure 19 - Receiver fault (relays wrong value)



majority(a1,a2,a3) = majority(v,0,v) = v, Receiver fault

Figure 20 - Receiver fault (doesn't relay a value)

20949P-000800US George B. F. Yee, Reg. No. 37,478 Telephone: 560-326-2400 Inventors: Richard Meyer et al. Title: Failover Processing in a Storage System Sheets of drawings 20 of 36

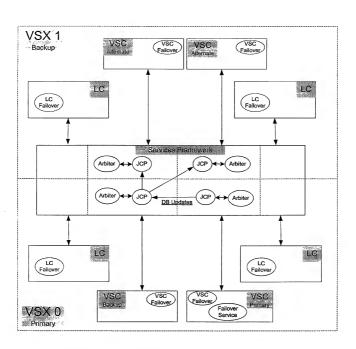


Figure 21A - Failover Service Architecture

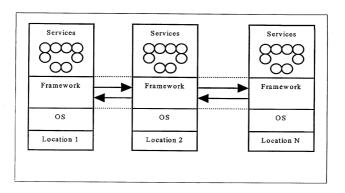


Figure 21B

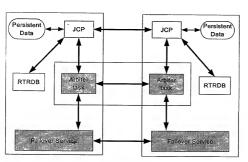


Figure 22 - An Arbiter for the Database

20949P-000800US George B. P. Yee, Reg. No. 37,478 Telephone: 650-326-2400 Inventors: Richard Meyer et al. Title: Failover Processing in a Storage System Sheets of drawings 22 of 3

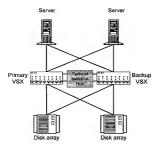


Figure 23 -Shared Link

Source (1)	Message Passing Interface (1)	SCSI Manager(1)	SCSI Manager (2)	Message Passing Interface (2)	Destination (2)
Send MSC to V	ifrom VSX (1) (SX (2) I/O	XFER One or moi	P Data Frames	fler for message	Message →
				1	l

Figure 24 - VSX to VSX Message Passing

20949P-000800US George B. F. Yee, Reg. No. 37,478 Telephone: 505-326-2400 Inventors: Richard Meyer et al. Title: Failover Processing in a Storage System Sheets of drawings 23 of 36

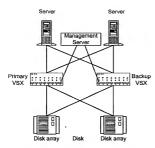


Figure 25 - Management Link

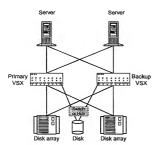


Figure 26 - Shared Disk

20949P-000800US George B. F. Yee, Reg. No. 37,478 Telephone: 509-326-2400 Inventors: Richard Meyer et al. Title: Failover Processing in a Storage System Sheets of drawings 24 of 36

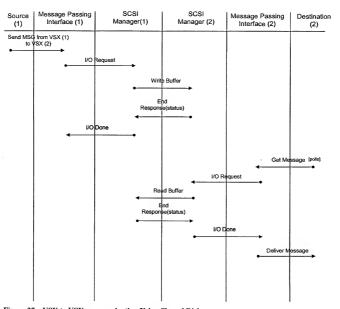
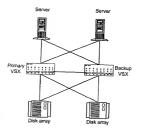


Figure 27 - VSX to VSX communication Using Shared Disk



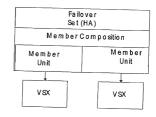
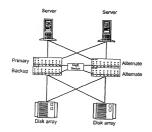


Figure 28 -2 Node HA Configuration



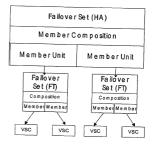


Figure 29 - Hierarchical HA Configuration

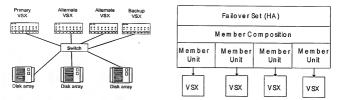


Figure 30 - N + 1 Nodes

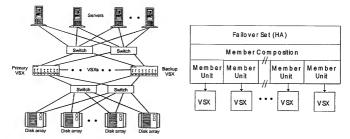
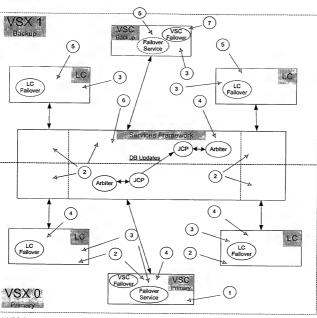


Figure 31 - N - Nodes

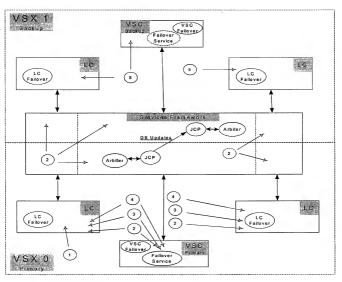
20949P-000800US George B. F. Yee, Reg. No. 37,478 Telephone. 650-326-2400 Inventors: Richard Meyer et al. Title: Failover Processing in a Storage System Sheets of drawings 27 of 36



- 1. VSC Crashes (Host Processor)
- 2. Rest of system detects VSC crash
- 3. Error Analysis determines Member fails, which translates into a "Primary Lost" event
- 4. Activate JCP in Master mode and enable the virtual services, Stop Ports on failed Primary
- 5. Reset affected devices, Cleanup reservations and locks. Set Unit Attention
- Restart management requests
 Restart RCON and FORMAT

Figure 32 - VSX Failover, Primary Fails

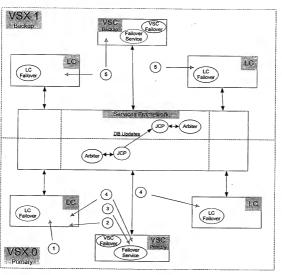
20949P-000800US George B. F. Yee, Reg. No. 37,478 Telephone: 650-326-2400 Inventors: Richard Meyer et al. Title: Failover Processing in a Storage System Sheets of drawings 28 of 36



- 1. LC Crashes (Host Processor)
- 2. Rest of system detects LC crash
- 3. Error Analysis determines IO Path fails for all devices (server and storage) on LC
- 4. Upstream hLUNs report CHECK CONDITION for all devices connected to ports on failed LC.
- RCON and FORMAT aborted, if necessary.
 5. Restart RCON and FORMAT, if necessary

Figure 33 - IO Path Failover - LC Fails

20949P-000800US George B. F. Yee, Reg. No. 37,478 Telephone: 650-326-2400 Inventors: Richard Meyer et al. Title: Failover Processing in a Storage System Sheets of drawings 29 of 36

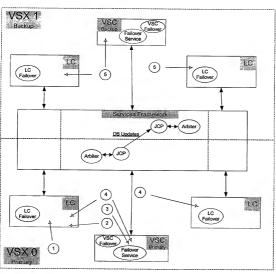


- FC ASIC Crashes
 LC detects FC ASIC crash
- 3. Error Analysis determines IO Path fails for all devices (server or storage) on FC ASIC
- 4. Upstream hLUNs report CHECK CONDITION for all devices connected to failed FC Ports. RCON
- and FORMAT aborted, if necessary

 5. Restart RCON and FORMAT, if necessary

Figure 34 - IO Path Failover - FC Port Fails

20949P-000800US George B. F. Yee, Reg. No. 37,478 Telephone: 650-326-2400 Inventors: Richard Meyer et al. Title: Failover Processing in a Storage System Sheets of drawings 30 of 36



- 1. Link down on port
- 2. LC detects FC Port link down
- Error Analysis determines IO Path fails for all devices (server or storage) on FC Port
 Upstream hLUNs report CHECK CONDITION for all devices connected to FC Port. RCON

i,

- and FORMAT aborted, if necessary

 5. Restart RCON and FORMAT, if necessary

Figure 35 - IO Path Failover - Link Down

20949P-000800US George B. F. Yee, Reg No 37,478 Telephone: 650-326-2400 Inventors: Richard Meyer et al. Title: Failover Processing in a Storage System Sheets of drawings 31 of 36

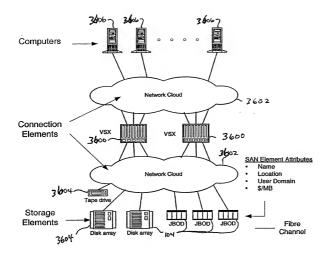


Fig. 36

20949P-000800US George B F. Yee, Reg. No. 37,478 Telephone: 650-326-2400 Inventors: Richard Meyer et al. Title: Failover Processing in a Storage System Sheets of drawings 32 of 36

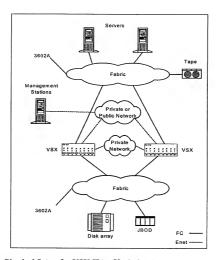


Figure 36 A Physical Setup for VSX-HA - Variation 1

20949P-000800US George B. F. Yee, Reg. No. 37,478 Telephone: 560-326-2400 Inventors: Richard Meyer et al. Title: Failover Processing in a Storage System Sheets of drawings 33 of 36

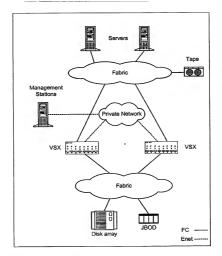


Figure 36 B Physical Setup for VSX-HA - Variation 2

20949P-000800US George B. F. Yee, Reg. No. 37,478 Telephone: 650-326-2400 Inventors: Richard Meyer et al. Title: Failover Processing in a Storage System Sheets of drawings 34 of 37.

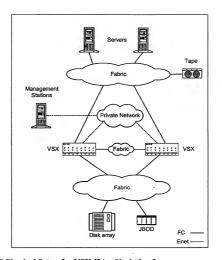


Figure 36 C Physical Setup for VSX-HA - Variation 3

20949P-000800US George B. F. Yee, Reg. No. 37,478 Telephone: 650-326-2400 Inventors: Richard Meyer et al. Title: Fallower Processing in a Storage System Sheets of drawings 35 of 36

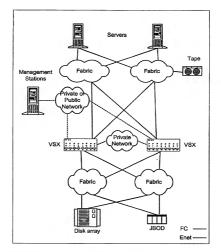


Figure 36 D Physical Setup for VSX-HA - Variation 4

20949P-000800US George B. F. Yee, Reg. No. 37,478 Telephone: 650-326-2400 Inventors: Richard Meyer et al. Title: Failover Processing in a Storage System Sheets of drawings 36 of 36

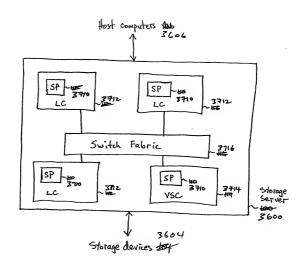


Fig. 37